ANL-IPG Collaborators Program-Level Overview

Lee Liming

Distributed Systems Laboratory

Argonne National Laboratory

the globus project™
www.globus.org

ANL Program-level Overview

- Application Frameworks
  - Bridging the gap between Grid applications and Grid infrastructure
- Software Usability
  - Improving usability for existing Grid software
- Technical Support
  - Helping IPG to use our Grid software
- Standardization
  - Forming/maintaining Grid community standards, and moving toward industry adoption

the globus project™
www.globus.org

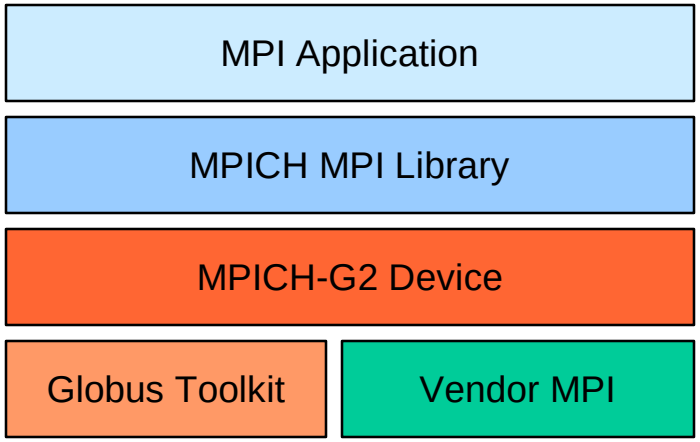
Application Frameworks

- Condor-G
  - Developed at the University of Wisconsin, replaces globus-job-* and globusrun tools as the premier client interface to the Grid
- MPICH-G2
  - Deploying Grid-enabled MPI
- Data Grid
  - Infrastructure technologies for data-intensive science
- Scientific Portals
  - Not much new in FY01, previous work is being used for cool new stuff.

the globus project™
www.globus.org

IPG MPICH-G2 Deployment

- MPICH-G2 is software developed in FY00 (supported by IPG/PACI funding) that allows MPI applications to run on a Grid.
- In FY01, MPICH-G2 was deployed on IPG resources by IPG Engineering staff, supported by ANL personnel.
- Available today on most IPG resources.



Globus Toolkit is used for WAN IPC. (Security, Seamless Access to Remote Resources)

Vendor MPI is used for LAN IPC. (High Performance, Architecture-optimized)

the globus project™
www.globus.org

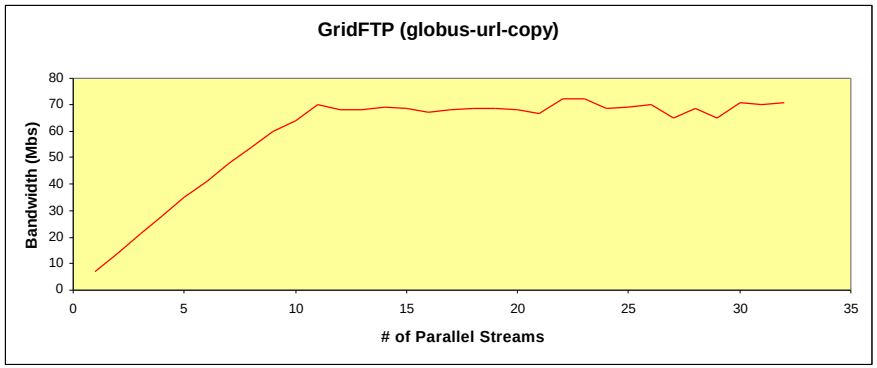
Data Grid Technologies

- The Globus Data Grid technologies address large-scale multi-institutional data management issues.
  - GridFTP provides a high-performance, secure, configurable data transport protocol for wide-area environments.
  - Replica Management provides a mechanism for managing large distributed data sets.

GridFTP for Efficient WAN Data Transfer

TODAY

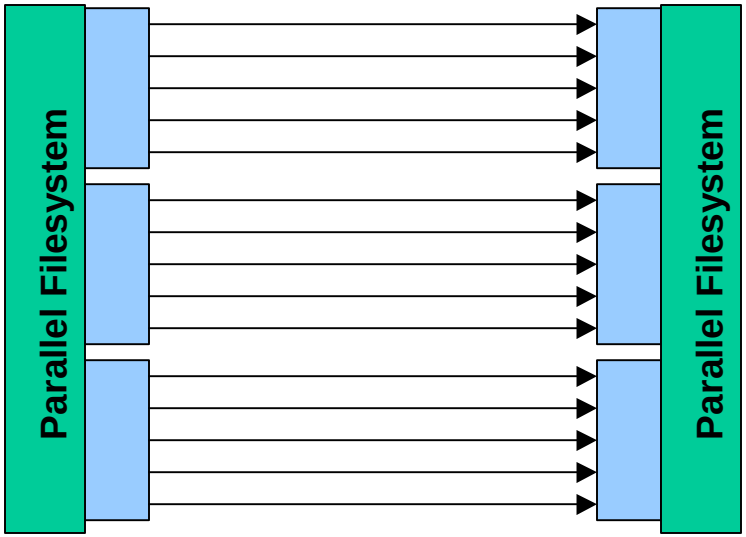
- Secure authentication
- Parallel transfer gets job done quickly
- Partial file access gets only required data
- Up to 2.8Gb/s using a striped server architecture



FUTURE

- Integrate striped GridFTP with HPSS data storage
  (work to be funded by ASCI)

Parallel Transfer
Fully utilizes bandwidth of network interface on single nodes.



Striped Transfer
Fully utilizes bandwidth of Gb+ WAN using multiple nodes.

the globus project™
www.globus.org

Data Grid Status

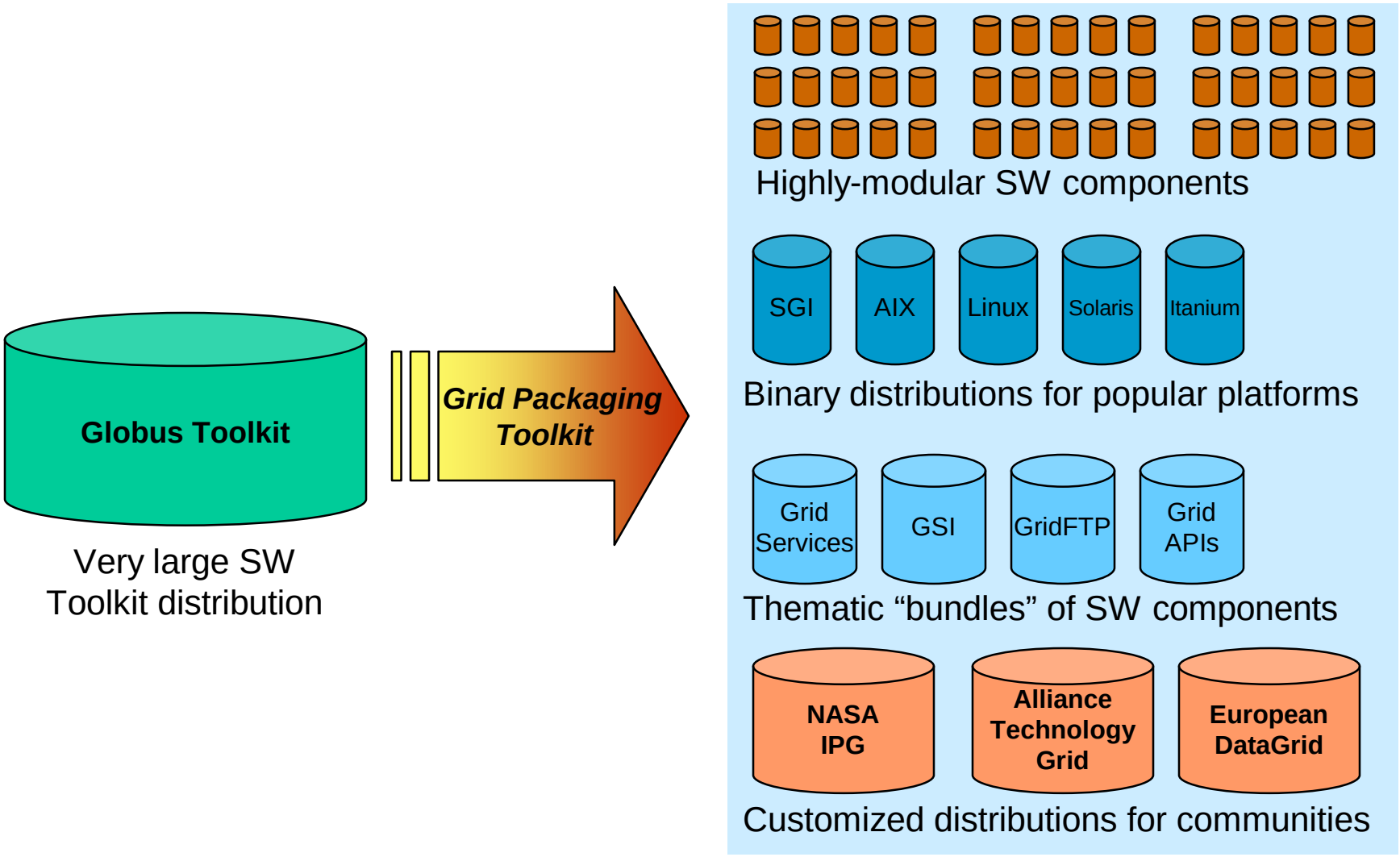
- Our Data Grid software was used by various Grid projects throughout FY01.
  - GridFTP tested and improved, applications developed.
  - Replica Catalog and Management APIs implemented and in use by Grid projects.
  - A completed version was released to the public in November.
- NASA has GSI-enabled FTP available today, and can move to GridFTP whenever needed.
  - Server program is very similar to GSI-FTP.
  - Advanced client and developer APIs provided with Globus Toolkit 2.0.

the globus project™
www.globus.org

Software Usability

- The Grid Packaging Toolkit (GPT) is a technology that manages the complexity of distributing and installing Grid software.
  - Challenges include: multiple platforms, compiler settings, version management, individual component installation, site-specific customization
- The GPT makes it easier for the scientific community to install and configure Grid software (such as the Globus Toolkit 2.0).

Sorting Out Grid Software



the globus project™
www.globus.org

Globus Toolkit Repackaging Status

- NCSA developed the GPT.
- ANL and ISI worked closely with NCSA to apply GPT to the Globus Toolkit 2.0.
  - Testing by external users began in August.
  - A completed version was released to the public in November.
- NASA is moving toward deployment of Globus Toolkit 2.0.
  - Testbed hardware was/is delayed?
  - Binary distributions may be helpful.
  - Could provide custom binary distribution (tools, APIs) for deployment to IPG user desktops?

the globus project™
www.globus.org

Technical Support

- The Globus Project provides technical support on Globus Toolkit software to NASA IPG team members as needed.
  - Discussions at IPG Engineering meetings
  - Email technical support
- Email support is primarily in three areas.
  - Help the IPG Engineering team to build the IPG infrastructure using Grid services and tools from the Globus Toolkit.
  - Help NASA application scientists to develop Grid-enabled software using Globus Toolkit APIs.
  - Help IPG Help Desk to support users.
    (Is this opportunity being utilized?)

the globus project™
www.globus.org

Technical Support Status

- Technical support requests from NASA were down considerably in FY01 compared to FY99 and FY00.
  - FY99: 54 requests
  - FY00: 117 requests
  - FY01: 15 requests
- Added an escalation capability for NASA support.

the globus project™
www.globus.org

Standardization

- The Global Grid Forum (GGF) promotes interoperability and best practices for Grid communities.
- Its goals are to reduce duplication of effort and to increase the quality of work in the Grid Computing field.
- Charlie Catlett (ANL) is the GGF Chair, Ian Foster (ANL) is on the steering committee, and Rick Stevens (ANL) is on the advisory committee.
- NCSA and ANL staff chair 6 of the 13 GGF working groups.

the globus project™
www.globus.org

Standardization

- GGF holds three meetings a year, alternating between the U.S. and Europe.
- GGF-2 (Washington D.C., July 15-18, 2001)
  - 340 Participants, 180 Organizations, 20 Countries
  - 36 documents
  - 48 working group sessions
  - 6 tutorials
  - 6 BOFs
  - 2 days General Updates

the globus project™
www.globus.org

Industry Adoption

- ANL and USC-ISI aggressively pursued industry adoption of our Grid software.
- On November 12th, major industry players (IBM, Compaq, Platform, NEC, etc.) endorsed the Globus Toolkit as a standard Grid platform for users of their products.
- Industry support for the Globus Toolkit is both financial and intellectual. We are working with partners in industry to evolve Grid technologies in directions that will ensure long-term support for Grid computing.